Mathematical Modelling, Vol. 7, pp. 507–523, 1986
Printed in the U.S.A. All rights reserved.
0270–0255/86 \$3.00 + .00
Copyright © 1986 Pergamon Journals Ltd.

APPLICATIONS OF NONSTANDARD ANALYSIS TO PARTIAL DIFFERENTIAL EQUATIONS—I. THE DIFFUSION EQUATION

FENG HANQIAO,† D. F. ST. MARY‡ and FRANK WATTENBERG

Department of Mathematics and Statistics
University of Massachusetts
Amherst, MA 01003

Communicated by Gian-Carlo Rota

(*Received February* 1986)

Abstract—This paper begins by developing a basis for using * finite difference equations to model physical phenomena. Under appropriate conditions the solution **F** of a * finite difference equation has S-continuous "finite difference derivatives" up to order r. In these circumstances we can show that the standard function °**F** is a C^r-function and satisfies the differential equation corresponding to the original finite difference equation. The second part of the paper illustrates these techniques by applying them to the heat equation. In particular, we obtain a very nice model of the heat equation with initial conditions corresponding to all the heat concentrated at a single point.

I. INTRODUCTION AND PRELIMINARIES

The techniques and perspective of nonstandard analysis provide a very natural approach to partial differential equations. This approach offers three main advantages. First, it allows us to derive many natural PDEs using the intuitive approaches of physicists and engineers with complete rigor. Second, proofs are very natural and intuitive. Finally, these techniques, especially when combined with recent work on nonstandard analysis and probability theory, are extremely powerful. There is every reason to expect that these techniques will eventually lead to new results.

This paper is divided into four parts. The second section develops the basic results needed to apply nonstandard analysis to PDEs. That section contains all the new mathematics in this paper. The aim of the remaining sections is to illustrate this approach in a particularly simple setting by studying the diffusion or heat equation. In Sec. III we derive the k-dimensional diffusion equation from an extremely natural and intuitive, yet completely rigorous model. This approach allows us to work with any initial conditions using very elementary methods. In Sec. IV we illustrate this by giving a completely elementary analysis of the one-dimensional diffusion equation with initial conditions corresponding to all the diffusing material being at a single point at time zero.

Subsequent papers will discuss the wave equation and the Fokker–Planck or Kolmogorov forward equation.

This paper will presuppose familiarity with only the most basic techniques of nonstandard analysis. We will work with higher-order nonstandard models but do not require any additional properties. (e.g. saturation). See any one of the standard introductions to Nonstandard Analysis for the necessary background (e.g. [2–4]).

---

† Department of Mathematics, Shaanxi Normal University, Xian, Shaanxi, People's Republic of China.
‡ Partially supported by NSF Grant No. RII-8115410-01.

We will be working with higher-order nonstandard models $^*\mathbf{R}^k$ of Euclidean k-space $\mathbf{R}^k$. We use the usual notation $\vec{x} \approx \vec{y}$ for $\vec{x}$ infinitely close to $\vec{y}$ and $\mathrm{St}(\vec{x})$ for the standard part of $\vec{x}$ if $\vec{x}$ is finite. $\mathbf{Z}$ will denote the set of integers and $\mathbf{N}$ the set of non-negative integers.

II. MACHINERY

In this section we develop some basic machinery for the application of nonstandard analysis to PDEs. Throughout this section Δt will denote a positive infinitesimal.

II.1. *Definition*

Suppose T is an internal subset of $^*\mathbf{R}^k$. We define $^\circ T = \{\mathrm{St}(\vec{t}) \mid \vec{t} \in T \text{ and } \vec{t} \text{ is finite}\}$.

II.2. *Definition*

Suppose either

$$T = \{k\Delta t \mid k \in {}^*\mathbf{N}\} = \{0, \Delta t, 2\Delta t, \ldots\}$$

or

$$T = \{k\Delta t \mid k \in {}^*\mathbf{Z}\} = \{\ldots, -\Delta t, 0, \Delta t, \ldots\}.$$

(Note that in the first case $^\circ T = [0, \infty)$ and in the second case $^\circ T = \mathbf{R}$). Suppose $F: T \to {}^*\mathbf{R}$ is an internal function. Then F is said to be *S-continuous* provided

(i) $F(0)$ is finite.

(ii) For every finite $s, t \in T$ $\quad s \approx t \Rightarrow F(s) \approx F(t)$.

This is a variation of a definition due to Robinson [3, p. 113]. Propositions II.3 and II.4 are well known and easy to prove. (Note: Both propositions require our assumption that F is internal.)

II.3. *Proposition*

Condition (ii) of II.2 is equivalent to:

(ii′) For every finite $t \in T$ and every *standard* $\epsilon > 0$ there is a *standard* $\delta > 0$ such that

$$\forall s \in T \quad |s - t| < \delta \Rightarrow |F(s) - F(t)| < \epsilon.$$

Proof. Omitted.

II.4. *Proposition and definition*

Suppose $F: T \to {}^*\mathbf{R}$ is S-continuous then

(i) For every finite $t \in T$ $F(t)$ is finite.

Hence, we can define a standard function $^\circ F: {}^\circ T \to \mathbf{R}$ as follows.

For $x \in {}^\circ T$ choose any $t \in T$ such that $x \approx t$. Define $^\circ F(x) = \mathrm{St}(F(t))$.

(ii) $^\circ F$ is well defined. [Since $t \approx x$, $t' \approx x \Rightarrow t \approx t'$ so $F(t) \approx F(t')$].

(iii) $^\circ F$ is continuous.

Proof. Omitted.

We will be working with the finite difference form of the derivative. This is defined in the obvious way.

II.5. *Definition*

Suppose $F: T \to {}^*\mathbf{R}$ is internal. Define $D_t F: T \to {}^*\mathbf{R}$ by

$$D_t F(t) = \frac{F(t + \Delta t) - F(t)}{\Delta t}.$$

The higher-order finite difference operators $D_t^{(k)} F(t)$ are defined inductively

$$D_t^{(1)} F(t) = D_t F(t),$$
$$D_t^{(k+1)} F(t) = D_t(D_t^{(k)} F)\,(t).$$

For example,

$$D_t^{(2)} F(t) = \frac{F(t + 2\,\Delta t) - 2F(t + \Delta t) + F(t)}{(\Delta t)^2}.$$

For standard functions the ordinary derivatives and the finite difference derivatives are related in the obvious way.

II.6. *Lemma*

Suppose $f: {}^\circ T \to \mathbf{R}$ is a standard C^r-function. Being moderately abusive of notation we write ${}^*f: T \to {}^*\mathbf{R}$ for ${}^*f\,|_T$. Then for $k = 1, 2, \ldots, r$ $D_t^{(k)} {}^*f$ is S-continuous and ${}^\circ(D_t^{(k)} {}^*f) = f^{(k)}$.

Proof:

(i) The fundamental theorem of calculus applies to any ${}^*C^1$ function. Hence,

$$\begin{aligned} D_t F(t) &= \frac{1}{\Delta t}\,[F(t + \Delta t) - F(t)] \\ &= \frac{1}{\Delta t} \int_t^{t+\Delta t} F'(s)\,\mathrm{d}s \\ &= \frac{1}{\Delta t} \int_0^{\Delta t} F'(t + s)\,\mathrm{d}s. \end{aligned}$$

(ii) We prove by induction

$$D_t^{(k)} {}^*f(t) = \frac{1}{(\Delta t)^k} \int_0^{\Delta t} \cdots \int_0^{\Delta t} {}^*f^{(k)}(t + s_1 + \cdots + s_k)\,\mathrm{d}s_1 \cdots \mathrm{d}s_k,$$

$k = 1$ follows from (i) above.

$$\begin{aligned}
D_t^{(k+1)*}f(t) &= D_t D_t^{(k)*}f(t) = \text{by induction} \\
&= D_t \frac{1}{(\Delta t)^k} \int \cdots \int {}^*f^{(k)}(t + s_1 + \cdots + s_k)\, \mathrm{d}s_1 \cdots \mathrm{d}s_k \\
&= \frac{1}{(\Delta t)^k} \int \cdots \int D_t{}^*f^{(k)}(t + s_1 + \cdots + s_k)\, \mathrm{d}s_1 \cdots \mathrm{d}s_k \\
&= \frac{1}{(\Delta t)^{k+1}} \int \cdots \int {}^*f^{(k+1)}(t + s_1 + \cdots + s_{k+1})\, \mathrm{d}s_1 \cdots \mathrm{d}s_{k+1}.
\end{aligned}$$

(iii) Now since ${}^*f^{(k+1)}$ is S-continuous and $\Delta t \approx 0$

$$\frac{1}{(\Delta t)^{k+1}} \int \cdots \int {}^*f^{(k+1)}(t + s_1 + \cdots + s_{k+1})\, \mathrm{d}s_1 \cdots \mathrm{d}s_{k+1} \approx {}^*f^{(k+1)}(t) \approx {}^*f^{(k+1)}(\mathrm{St}(t))$$

which completes the proof.

The following lemma is the key lemma for our work.

II.7. *Key lemma*

Suppose $F: T \to {}^*\mathbf{R}$ and
(i) $F(0)$ is finite.
(ii) $D_t F$ is S-continuous.

Then
(i) F is S-continuous.
(ii) ${}^\circ F$ is continuously differentiable and $\forall x \in {}^\circ T$ $({}^\circ F)'(x) = {}^\circ(D_t F)(x)$.

Proof:

(i) Suppose $s, t \in T$ are finite and $s \approx t$. Without loss of generality $s < t$.

Let $s = m\Delta t, \qquad t = n\Delta t.$

Let $M = \max\{\, | D_t F(i\Delta t) | \mid m \le i < n\}.$

M is finite.

$$\begin{aligned}
F(t) - F(s) &= \sum_{i=m}^{n-1} [F((i+1)\Delta t) - F(i\Delta t)] \\
&= \sum_{i=m}^{n-1} \Delta t D_t F(i\Delta t).
\end{aligned}$$

So

$$\begin{aligned}
| F(t) - F(s) | &\le \sum_{i=m}^{n-1} \Delta t \,| D_t F(i\Delta t) | \\
&\le \sum_{i=m}^{n-1} \Delta t M \\
&\le M((n-1)\Delta t - m\Delta t) \approx 0
\end{aligned}$$

since $n\Delta t \approx m\Delta t$ and M is finite.

(ii) By Proposition II.4 $°(D_tF)$ is continuous so (ii) will follow if we can show

$$\lim_{h\to 0} \frac{°F(x+h) - °F(x)}{h} = °(D_tF)(x).$$

Suppose $\epsilon > 0$ is standard. Choose $t = m\Delta t \in T$ such that $t \approx x$. Since D_tF is S-continuous, Proposition II.3 yields a standard $\delta > 0$ such that

$$\forall s \in T \qquad |s - t| < \delta \Rightarrow |D_tF(s) - D_tF(t)| < \epsilon/2.$$

Now suppose $y \in °T$ and $0 < |x - y| < \delta/2$. Choose $s = n\Delta t \in T$ such that $s \approx y$. Without loss of generality $t < s$

$$F(s) - F(t) = \sum_{i=m}^{n-1} \Delta t D_tF(i\Delta t)$$

$$(*) \qquad \frac{F(s) - F(t)}{(n-m)\Delta t} = \sum_{i=m}^{n-1} D_tF(i\Delta t)/(n-m).$$

Since $(n - m)\Delta t \not\approx 0$ the left-hand side of $(*)$ is infinitely close to

$$\frac{F(y) - F(x)}{y - x}.$$

But for $i = m, \ldots, n - 1$

$$D_tF(t) - \frac{\epsilon}{2} < D_tF(i\Delta t) < D_tF(t) + \frac{\epsilon}{2}.$$

So the right-hand side of $(*)$ satisfies

$$D_tF(t) - \frac{\epsilon}{2} \le \sum_{i=m}^{n-1} D_tF(i\Delta t)/(n-m) \le D_tF(t) + \frac{\epsilon}{2}.$$

So

$$\left| \frac{F(y) - F(x)}{y - x} - °D_tF(t) \right| < \epsilon$$

which completes the proof.

II.8. *Corollary*

Suppose $F: T \to {}^*\mathbf{R}$ is internal and
(i) $F(0), D_tF(0), \ldots, D_t^{(r-1)}F(0)$ are finite.
(ii) D_t^rF is S-continuous.
Then $°F$ is C^r and for $k = 1, 2, \ldots, r$

$$(°F)^{(k)}(x) = °(D_t^{(k)})(x).$$

Proof:

Immediate.

II.9. *Corollary*

Suppose $F: T \to {}^*\mathbf{R}$ is internal and
(i) $F(0), D_tF(0), \ldots, D_t^{(r)}F(0)$ are finite.
(ii) For every finite $t \in T$ $\quad D_t^{(r+1)}F(t)$ is finite.
Then ${}^\circ F$ is C^r and for $k = 1, 2, \ldots, r$

$$({}^\circ F)^{(k)}(x) = {}^\circ(D_t^{(k)})(x).$$

Proof:

This follows immediately from a straightforward verification of the fact that (ii) implies that $D_t^{(r)}F$ is S-continuous.

In our treatment of the diffusion equation we will be working with k-space dimensions and one time dimension. Everything generalizes in the obvious way.

II.10. *Definition*

Suppose k is a positive integer and Δx and Δt are positive infinitesimals. Let

$$\Lambda^t = \{(i_1\Delta x, \ldots, i_k\Delta x) \mid i_1, \ldots, i_k \in {}^*\mathbf{Z}\},$$
$$T = \{i\Delta t \mid i \in {}^*\mathbf{N}\}.$$

Suppose $F: \Lambda^k \times T \to {}^*\mathbf{R}$ is internal. F is *S-continuous* iff
(i) $F(\vec{0}, 0)$ is finite.
(ii) For every finite $(\vec{x}, t)$

$$(\vec{y}, s) \approx (\vec{x}, t) \Rightarrow F(\vec{y}, s) \approx F(\vec{x}, t).$$

Let $\vec{e}_1, \ldots, \vec{e}_k$ be the usual basis for $\mathbf{R}^k$. For $i = 1, 2, \ldots, k$ define

$$D_{x_i}F(\vec{x}, t) = \frac{F(\vec{x} + \Delta x\vec{e}_i, t) - F(\vec{x}, t)}{\Delta x}$$

and

$$D_tF(\vec{x}, t) = \frac{F(\vec{x}, t + \Delta t) - F(\vec{x}, t)}{\Delta t}.$$

All our results in this section generalize easily to $F: \Lambda^k \times T \to {}^*\mathbf{R}$ and ${}^\circ F: \mathbf{R}^k \times [0, \infty) \to \mathbf{R}$. The proofs of Lemma II.6 and Corollaries II.8 and II.9 apply to the multidimensional case using well-known standard results. (See, e.g. [1, Theorem 6–18]).

The proof of the following proposition is completely straightforward.

II.11. *Proposition*

Suppose $F: \Lambda^k \to {}^*\mathbf{R}$ is S-continuous and E is a standard open subset of $\mathbf{R}^k$ then

$$\int_E {}^\circ F \, dx_1 \cdots dx_k = \mathrm{St}\Big(\sum_{\vec{x} \in {}^*E \cap \Lambda^k} F(\vec{x})(\Delta x)^k\Big).$$

We now turn to the diffusion equation.

III. A MODEL OF DIFFUSION

Our model of diffusion involves a quantity of "ink" diffusing or jumping around on the k-dimensional grid Λ^k from Sec. II. We begin with one unit of ink placed on the grid according to some *initial conditions*. Formally, we are given an internal function $P:\Lambda^k \to {}^*[0, 1]$. $P(\vec{x})$ determines the amount of ink initially placed at $\vec{x}$. Since we work with one unit of ink

$$\sum_{\vec{x}\in\Lambda^k} P(\vec{x}) = 1.$$

The ink jumps around the grid according to the following rules.

(1) The ith jump occurs between time $(i - 1)\Delta t$ and $i\Delta t$.

(2) At each jump the "blob" of ink at each grid point $\vec{x}$ breaks into $(2k + 1)$ smaller blobs.

 (a) For $i = 1, 2, \ldots, k$ a fraction p_i of the blob at x jumps to $\vec{x} + \Delta x\vec{e}_i$.

 (b) For $i = 1, 2, \ldots, k$ a fraction p_i of the blob at $\vec{x}$ jumps to $x - \Delta x\vec{e}_i$.

 (c) A fraction $(1 - 2\Sigma p_i)$ of the blob at $\vec{x}$ remains at $\vec{x}$,

where $0 < p_i \le \frac{1}{2}$ and $\sum_1^k p_i \le \frac{1}{2}$.

Our formalization of this model is straightforward.

Let $C = \{c_0, c_1, \ldots, c_{2k}\}$. Think of C as the set of faces of a $(2k + 1) -$ sided die which is weighted so that the probability of each face coming up on a random roll is given by

(i) The probability of c_0 is $1 - 2\sum_{i=1}^k p_i$. The face c_0 "instructs" an ink molecule to remain at its current grid point.

(ii) the probability of c_{2i-1} is p_i for $i = 1, 2, \ldots, k$. The face c_{2i-1} "instructs" an ink molecule to jump from its current location at grid point $\vec{x}$ to $\vec{x} + \Delta x\vec{e}_i$.

(iii) The probability of c_{2i} is p_i for $i = 1, 2, \ldots, k$. The face c_{2i} "instructs" an ink molecule to jump from its current location at grid point $\vec{x}$ to $\vec{x} - \Delta x\vec{e}_i$.

III.1. *Definition*

Let R be an infinite, positive * integer such that $R\Delta t$ is infinite.

Let $\Omega = \Lambda^k \times C^R$.

Think of Ω as being a set of very tiny blobs of ink.

A typical element $\vec{\omega}$ of Ω looks like

$$\vec{\omega} = (\omega_0, \omega_1, \omega_2, \ldots),$$

where $\omega_0 \in \Lambda^k$, $\omega_1, \omega_2, \ldots \in C$. The "size" of each blob is determined by

(i) $l_0(\vec{\omega}) = P(\omega_0)$, the size of the initial blob at the grid point ω_0.

(ii) For $j \ge 1$

$$\begin{aligned} l_j(\vec{\omega}) &= p_i && \text{if } \omega_j = c_{2i-1} \text{ or } c_{2i} \\ &= 1 - 2\Sigma p_i && \text{if } \omega_j = c_0 \end{aligned}$$

and

$$S(\vec{\omega}) = \prod_{j=0}^{R} l_j(\vec{\omega}).$$

ω_0 determines the initial location of $\vec{\omega}$ and $\omega_1, \omega_2, \ldots, \omega_R$ describe how $\vec{\omega}$ jumps around.

More formally, we define

$$X:\Omega \times T \to \Lambda^k$$

where

$$T = \{0, \Delta t, \ldots, R\Delta t\}$$

by

$$X(\vec{\omega}, 0) = \omega_0$$

$$X(\vec{\omega}, j\Delta t) = \begin{cases} X(\vec{\omega}, (j-1)\Delta t) & \text{if } \omega_j = c_0 \\ X(\vec{\omega}, (j-1)\Delta t) + \Delta x\vec{e}_i & \text{if } \omega_j = c_{2i-1} \\ X(\vec{\omega}, (j-1)\Delta t) - \Delta x\vec{e}_i & \text{if } \omega_j = c_{2i}. \end{cases}$$

At any given instant of time $t = j\Delta t \in T$ there is a blob of ink at a grid point $\vec{x}$ which is made up of many blobs from Ω. The size of this composite blob is given by the sum

$$\sum_{X(\vec{\omega},t)=\vec{x}} S(\vec{\omega}).$$

We will keep track of the amount of ink at each point $\vec{x}$ at each time t by a function

$$F:\Lambda^k \times T \to {}^*[0, \infty)$$

defined by

$$F(\vec{x}, t) = \sum_{X(\vec{\omega},t)=\vec{x}} S(\vec{\omega}) \cdot \frac{1}{(\Delta x)^k}.$$

Thus, $F(\vec{x}, t)\,(\Delta x)^k$ is the fraction of the total quantity of ink which is at $\vec{x}$ at time t.

In particular, if $E \subseteq \Lambda^k$ the total fraction of ink in E at time t is

$$\sum_{\vec{x}\in E} F(\vec{x}, t)\,(\Delta x)^k.$$

Thus, $F(\vec{x}, t)$ is analogous to the usual standard density function $\phi(\vec{x}, t)$.

We will use the notation $H(\vec{x})$ for the initial condition, $H(\vec{x}) = F(\vec{x}, 0) = P(\vec{x})/(\Delta x)^k$. One particular initial condition to have in mind is

$$H(\vec{x}) = \begin{cases} \dfrac{1}{(\Delta x)^k} & \vec{x} = \vec{0} \\ \\ 0 & \vec{x} \neq \vec{0}. \end{cases}$$

This corresponds to all the ink starting at $\vec{0}$ at time 0.

$F(\vec{x}, t)$ evolves according to the finite difference initial value problem.

$$\begin{aligned} F(\vec{x}, 0) &= H(\vec{x}) \\ F(\vec{x}, t + \Delta t) &= (1 - 2\Sigma p_i)F(\vec{x}, t) \\ &\quad + \Sigma p_i[F(\vec{x} + \Delta x\vec{e}_i, t) + F(\vec{x} - \Delta x\vec{e}_i, t)]. \end{aligned} \tag{III.1}$$

So

$$F(\vec{x}, t + \Delta t) - F(\vec{x}, t) = \Sigma p_i[F(\vec{x} + \Delta x\vec{e}_i, t) - 2F(\vec{x}, t) + F(\vec{x} - \Delta x\vec{e}_i, t)]$$

and

$$\frac{F(\vec{x}, t + \Delta t) - F(\vec{x}, t)}{\Delta t} = \Sigma p_i \frac{(\Delta x)^2}{\Delta t}\left[\frac{F(\vec{x} + \Delta x\vec{e}_i, t) - 2F(\vec{x}, t) + F(\vec{x} - \Delta x\vec{e}_i, t)}{(\Delta x)^2}\right]$$

or,

$$D_t F(\vec{x}, t) = \Sigma p_i \frac{(\Delta x)^2}{\Delta t} D^2_{x_i x_i} F(\vec{x} - \Delta x\vec{e}_i, t). \tag{III.2}$$

Now, we add the additional assumption that $p_i\,(\Delta x)^2/\Delta t$ is finite for $i = 1, 2, \ldots, k$ and let $\alpha_i = \mathrm{St}(p_i\,(\Delta x)^2/\Delta t)$

Notice that (III.2) is suggestive of the usual diffusion or heat equation

$$\frac{\partial\phi}{\partial t} = \Sigma\alpha_i \frac{\partial^2\phi}{\partial x_i^2} \tag{III.3}$$

with $\phi = {}^{\circ}F$.

The remainder of this section is devoted to justifying the transition from (III.2) to (III.3).

III.2. *Lemma*

Suppose $G:\Lambda^k \times T \to {}^*[0, \infty)$ satisfies Eq. (III.1). That is,

$$G(\vec{x}, t + \Delta t) = (1 - 2\Sigma p_i)G(\vec{x}, t) + \Sigma p_i[G(\vec{x} + \Delta x\vec{e}_i, t) + G(\vec{x} - \Delta\vec{x}e_i, t)]$$

and

$$\forall \vec{x} \in \Lambda^k \qquad 0 \le G(\vec{x}, 0) \le M.$$

Then

$$\forall \vec{x} \in \Omega^k \qquad t \in T \qquad 0 \le G(\vec{x}, t) \le M.$$

Proof.

Straightforward.

III.3. *Lemma*

Suppose $G:\Lambda^k \times T \to {}^*[0, \infty)$ satisfies Eq. (III.1). Then so does $D_{x_i}G$.

Proof.

Straightforward.

III.4. *Theorem*

Suppose F is as above, M is a standard positive real, $t_0 \in T$, and the function $G(x) = F(x, t_0)$ satisfies

$$\forall \vec{x} \in \Lambda^k \qquad \forall r \leq 4 \qquad \forall i_1, \ldots, i_r$$
$$| D^{(r)}_{x_{i_1} \cdots x_{i_r}} G(\vec{x}) | \leq M.$$

Then on $\Lambda^k \times T \cap {}^*[t_0, \infty)$

$$D_t F,\ D_{x_i} F,\ D^{(2)}_{x_i x_j} F$$

are all S-continuous and bounded by M. Therefore, F is S-continuous and, denoting ${}^\circ F$ by ϕ,

$$\frac{\partial \phi}{\partial t} = {}^\circ(D_t F)$$
$$\frac{\partial^2 \phi}{\partial x_i^2} = {}^\circ(D^2_{x_i x_i} F).$$

Therefore on $\mathbf{R}^k \times [\mathrm{St}(t_0), \infty)$ (III.3) holds by taking the standard part of both sides of (III.2).

Proof.

By Lemma III.3 for $0 \leq r \leq 4$ and all $i_1, \ldots, i_r$ $D^{(r)}_{x_{i_1}, \ldots, x_{i_r}} F$ satisfies Eq. (III.1). Therefore by Lemma III.2 for all $\vec{x} \in \Lambda^k$ and $t \in T$, $t \geq t_0$

$$| D^{(r)}_{x_{i_1}, \ldots, x_{i_r}} F(\vec{x}, t) | \leq M.$$

Now, for $\vec{x} \in \Lambda^k$, $t \geq t_0$

$$D_t F(x, t) = \sum_{i=1}^{k} p_i \frac{(\Delta x)^2}{\Delta t} D^{(2)}_{x_i x_i}$$

is bounded by

$$\sum_{i=1}^{k} p_i \frac{(\Delta x)^2}{\Delta t} M$$

which is finite.

Similarly $D^{(2)}_{tt} F(\vec{x}, t)$ has a finite bound as do $D^{(2)}_{t x_i} F(\vec{x}, t)$, $D^{(3)}_{x_i x_j x_k} F(\vec{x}, t)$ and $D^{(3)}_{t x_i x_j} F(\vec{x}, t)$ for $\vec{x} \in \Lambda^k$, $t \geq t_0$.

The result follows immediately from the multivariable version of Corollary II.9.

III.5. *Corollary*

Suppose F, H are as above, M is a standard positive number and H satisfies

$$\forall x \in \Lambda^k \qquad \forall r \leq 4 \qquad i_1, \ldots, i_r$$
$$| D^r_{i_1, \ldots, i_r} H(\vec{x}) | \leq M.$$

Then H and F are S-continuous. Let $h = {}^\circ H$, $\phi = {}^\circ F$. Then

$$\frac{\partial \phi}{\partial t} = \Sigma \alpha_i \frac{\partial^2 \phi}{\partial x_i^2}$$

$$\phi(x, 0) = h(x).$$

In particular, by the multivariable version of Lemma II.6 we can apply Corollary III.5 to standard initial conditions. The only difficulty is a very minor one. If we let $H(\vec{x}) = {}^*h(\vec{x})$ then $\Sigma H(x)$ will be infinitely close to but not necessarily equal to 1. So let

$$H(\vec{x}) = \frac{{}^*h(\vec{x})}{\Sigma {}^*h(\vec{x})}.$$

Part of the power of this approach is its ability to handle initial conditions which are very irregular and, in particular, do not satisfy the hypotheses of Corollary III.5. The next section illustrates this power.

IV. AN ELEMENTARY TREATMENT OF A SPECIAL CASE

In this section we will give a completely elementary treatment of the special case $k = 1$, $p_i = \frac{1}{2}$ with the initial condition

$$H(x) = \begin{cases} \dfrac{1}{\Delta x} & x = 0 \\ 0 & x \neq 0. \end{cases}$$

Notice that H is not S-continuous. For simplicity we will assume $\Delta t = (\Delta x)^2$.

The function $F(x, t)$ will not be S-continuous even for t not infinitely close to 0 since $F(j\Delta x, i\Delta t)$ is nonzero only for $i \equiv j$ (Mod 2).

In this situation the treatment in Sec. III can be greatly simplified.

IV.1. *Definition*

Let $\Omega_0 = \{-1, +1\}^R$ so that $\vec{\omega} \in \Omega_0$ looks like:

$$\vec{\omega} = (\omega_1, \omega_2, \ldots, \omega_R),$$

where each ω_i is $+1$ or -1. Let $X{:}\Omega_0 \times T \to \Lambda^1$ be given by

$$X(\vec{\omega}, 0) = 0$$

$$X(\vec{\omega}, j\Delta t) = X(\vec{\omega}, (j-1)\Delta t) + \Delta x \omega_j \quad 1 \le j \le R$$

and

$$F{:}\Lambda^1 \times T \to {}^*[0, \infty)$$

$$F(x, t) = \frac{\mathrm{Card}\{\vec{\omega} \in \Omega_0 \mid X(\vec{\omega}, t) = x\}}{2^R \cdot \Delta x}.$$

In particular, notice

$$F(0, 0) = \frac{\mathrm{Card}(\Omega_0)}{2^R \cdot \Delta x} = \frac{1}{\Delta x} = H(0)$$

and,

$$F(x, t + \Delta t) = \tfrac{1}{2}[F(x + \Delta x, t) + F(x - \Delta x, t)].$$

The first step is to look at $F(x, t)$ for $t \approx 0$. The calculation is an absolutely straightforward one involving variance.

IV.2. *Definition*

Suppose $F: \Lambda^1 \times T \to {}^*[0, \infty)$ is as above. Define $\mathrm{Var}(F, t)$ by

$$\mathrm{Var}(F, t) = \sum_{x \in \Lambda^1} x^2(F(x, t)).$$

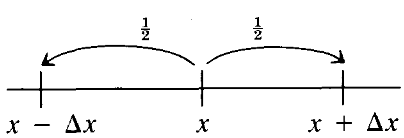
[Note: This is not quite the standard variance which would be $\sum x^2 F(x, t)\Delta x$.]

IV.3. *Lemma*

$$\mathrm{Var}(F, t + \Delta t) - \mathrm{Var}(F, t) = \Delta x.$$

Proof.

Look at the contribution of the ink at x at time t



The contribution of this ink to

$$\mathrm{Var}(F, t) \text{ is } x^2 F(x, t).$$

Its contribution to $\mathrm{Var}(F, t + \Delta t)$ is

$$\tfrac{1}{2}(x + \Delta x)^2 F(x, t) + \tfrac{1}{2}(x - \Delta x)^2 F(x, t)$$
$$= [\tfrac{1}{2}(x^2 + 2x\Delta x + (\Delta x)^2) + \tfrac{1}{2}(x^2 - 2x\Delta x + (\Delta x)^2)]F(x, t) = [x^2 + (\Delta x)^2]F(x, t).$$

The difference is $(\Delta x)^2 F(x, t)$ and summing over x we get

$$\mathrm{Var}(F, t + \Delta t) - \mathrm{Var}(F, t) = \sum_x (\Delta x)^2 F(x, t)$$
$$= (\Delta x)^2 \frac{1}{\Delta x} = \Delta x.$$

IV.4. *Corollary*

$$\operatorname{Var}(F, n\Delta t) = n\Delta x.$$

Proof.

Immediate.

IV.5. *Corollary*

Suppose $t = n\Delta t$ is infinitesimal. Then there is an infinitesimal α such that

$$\sum_{|x| \geq \alpha} F(x, t)\Delta x \approx 0.$$

In other words, for $t \approx 0$ almost all of the ink is infinitely close to 0.

Proof.

$$\sum_{|x| \geq \alpha} F(x, n\Delta t)\Delta x \leq \frac{1}{\alpha^2} \sum_{|x| \geq \alpha} x^2 F(x, n\Delta t)\Delta x = \frac{\Delta x}{\alpha^2} \sum_{|x| \geq \alpha} x^2 F(x, n\Delta t)$$

$$\leq \frac{\Delta x}{\alpha^2} \sum x^2 F(x, n\Delta t) = \frac{\Delta x}{\alpha^2} n\Delta x = \frac{n(\Delta x)^2}{\alpha^2} = \frac{n\Delta t}{\alpha^2}.$$

By assumption $n\Delta t \approx 0$. So if $\alpha = (n\Delta t)^{1/4}$ then $\alpha \approx 0$ and $n\Delta t/\alpha^2 = (n\Delta t)^{1/2} \approx 0$ and $\sum_{|x| \geq \alpha} F(x, n\Delta t)\Delta x \approx 0$ completing the proof.

Next we want to compute $F(x, t)$ for $x = 2k\Delta x$ finite and $t = 2n\Delta t$ and $t \not\approx 0$. Notice that $F((2k + 1)\Delta x, 2n\Delta t) = 0$.

A particular ink blob $\vec{\omega} = (\omega_1, \ldots, \omega_{2n}, \ldots, \omega_r)$ is at $2k\Delta x$ at time $2n\Delta t$ if and only if

$$\operatorname{Card}(\{i \mid 1 \leq i \leq 2n, \omega_i = +1\}) = n + k$$

and

$$\operatorname{Card}(\{i \mid 1 \leq i \leq 2n, \omega_i = -1\}) = n - k$$

Therefore,

$$F(2k\Delta x, 2n\Delta t) = \frac{1}{\Delta x} \frac{(2n)!}{(2)^{2n}(n + k)!(n - k)!}.$$

We need to apply Stirling's formula in the following form

IV.6. *Lemma*

Suppose ν is an infinite integer then $\nu! = e^{-\nu}\nu^{\nu}\sqrt{2\pi\nu}\,\delta_\nu$ where $\delta_\nu \approx 1$.

Proof.

Immediate consequence of Stirling's formula.

IV.7. *Lemma*

Suppose $n\Delta t \not\approx 0$ and $k\Delta x$ is finite. then $(n + k)$ and $(n - k)$ are infinite.

Proof.

We may assume $k \geq 0$.
Therefore, $n + k \geq n$ and since $n\Delta t \not\approx 0$, n is infinite so $(n + k)$ is also.
Now we look at $(n - k)$

$$k\Delta x \text{ is finite.}$$

Therefore, $k\Delta t = k(\Delta x)^2$ is infinitesimal.

So $$(n - k)\Delta t \approx n\Delta t \not\approx 0.$$

So $$n - k \text{ must be infinite.}$$

IV.8. *Definition*

Suppose $s, t \in {}^*\mathbf{R}$. We write $s \doteq t$ for $s/t \approx 1$. If s and t are finite and not infinitesimal then

$$s \doteq t \Leftrightarrow s \approx t.$$

Notice

(i) $s \doteq s'$ and $t \doteq t' \Rightarrow st \doteq s't'$

(ii) $s \doteq s' \Rightarrow \dfrac{1}{s} \doteq \dfrac{1}{s'}$

(iii) $s \doteq s'$ and $t \doteq t' \Rightarrow \dfrac{s}{t} \doteq \dfrac{s'}{t'}$.

By IV.6 and IV.7 we have

$$(2n)! \doteq e^{-2n}(2n)^{(2n)}\sqrt{4\pi n}$$
$$(n + k)! \doteq e^{-(n+k)}(n + k)^{(n+k)}\sqrt{2\pi(n + k)}$$
$$(n - k)! \doteq e^{-(n-k)}(n - k)^{(n-k)}\sqrt{2\pi(n - k)} .$$

IV.9. *Calculation*

We need to compute the right hand side of IV.1.

$$\frac{1}{\Delta x}\frac{(2n)!}{2^{2n}(n + k)!(n - k)!}$$
$$\doteq \frac{1}{\Delta x}\frac{e^{-2n}(2n)^{2n}\sqrt{4\pi n}}{2^{2n}e^{-(n+k)}(n + k)^{(n+k)}\sqrt{2\pi(n + k)}e^{-(n-k)}(n - k)^{(n-k)}\sqrt{2\pi(n - k)}}$$
$$= \frac{1}{\Delta x}\frac{n^{2n}\sqrt{4\pi n}}{(n + k)^{(n+k)}(n - k)^{(n-k)}2\pi\sqrt{n^2 - k^2}}$$
$$= \frac{1}{\Delta x}\frac{n^{2n}}{(n + k)^n(n - k)^n}\cdot\left(\frac{n - k}{n + k}\right)^k\cdot\frac{\sqrt{\pi n}}{\pi\sqrt{n^2 - k^2}} .$$

Let

$$(1) = \frac{n^{2n}}{(n+k)^n(n-k)^n}$$

$$(2) = \left(\frac{n-k}{n+k}\right)^k$$

$$(3) = \frac{\sqrt{\pi n}}{\pi\sqrt{n^2-k^2}\,\Delta x}.$$

(1): $$\frac{n^{2n}}{(n+k)^n(n-k)^n} = \frac{1}{(1+k/n)^n(1-k/n)^n} = \frac{1}{\left(1-k^2/n^2\right)^n} = \frac{1}{\left(1-\frac{k^2/n}{n}\right)^n}.$$

But

$$\frac{k^2}{n} = \frac{(k\Delta x)^2}{n\Delta t} = \frac{(x/2)^2}{t/2} = \frac{x^2}{2t}.$$

So $(1) \doteq e^{x^2/2t}$.

(2): $$\left(\frac{n-k}{n+k}\right)^k = \left(\frac{n+k-2k}{n+k}\right)^k = \left(1-\frac{2k}{n+k}\right)^k = \left(1-\frac{2k^2/(n+k)}{k}\right)^k.$$

But

$$\frac{2k^2}{n+k} = \frac{2k^2 2(\Delta x)^2}{(n+k)2(\Delta x)^2} = \frac{(2k\Delta x)^2}{2(n+k)(\Delta x)^2}.$$

But $$k(\Delta x)^2 \approx 0$$

So $$2(n+k)(\Delta x)^2 \approx 2n(\Delta x)^2 = 2n\Delta t = t.$$

So $(2) \doteq e^{-x^2/t}$.

(3): $$\frac{\sqrt{\pi n}}{\Delta x\pi\sqrt{n^2-k^2}} = \frac{1}{\sqrt{\pi}}\sqrt{\frac{n}{(n\Delta x)^2-(k\Delta x)^2}} = \frac{1}{\sqrt{\pi}}\sqrt{\frac{1}{n(\Delta x)^2-(k\Delta x)^2/n}}.$$

But $n(\Delta x)^2 = n\Delta t = t/2 \not\approx 0$, $(k\Delta x)^2$ is finite and n is infinite.

So $(k\Delta x)^2/n \approx 0$.

Therefore,

$$(3) \doteq \frac{1}{\sqrt{\pi}}\sqrt{\frac{2}{t}} = \frac{\sqrt{2}}{\sqrt{\pi t}}.$$

Putting this all together

$$F(x,t) \doteq \frac{\sqrt{2}}{\sqrt{\pi t}}e^{x^2/2t}e^{-x^2/t} = \frac{\sqrt{2}}{\sqrt{\pi t}}e^{-x^2/2t}.$$

IV.10. *Corollary*

Suppose $t = 2n\Delta t \not\approx 0$ and $a, b \in \mathbf{R}$ $a < b$. Then

$$\begin{aligned}\sum_{x\in{}^*(a,b)} F(x,t)\Delta x &= \sum_{\substack{x\in{}^*(a,b)\\ x=2k\Delta x}} F(x,t)\,\Delta x\\ &= \tfrac{1}{2}\sum_{\substack{x\in{}^*(a,b)\\ x=2k\Delta x}} F(x,t)(2\Delta x)\\ &\approx \int_a^b \frac{1}{\sqrt{2\pi t}}\,\mathrm{e}^{-x^2/2t}\,\mathrm{d}x.\end{aligned}$$

This is exactly the usual standard result.

This result allows us to handle any initial conditions via the following:

IV.11. *Corollary*

Suppose

$$F(x,0) = \begin{cases} a & x = x_0\\ 0 & x \neq x_0\end{cases}$$

and

$$F(x, t + \Delta t) = \tfrac{1}{2}[F(x + \Delta x, t) + F(x - \Delta x, t)]$$

and

$$t = 2n\Delta t \not\approx 0.$$

Then for $2k\Delta x$ finite

(i) $F(x_0 + (2k + 1)\Delta x, t) = 0$

(ii) $F(x_0 + 2k\Delta x, t) \doteq \dfrac{\sqrt{2}}{\sqrt{\pi t}}\,\mathrm{e}^{-(2k\Delta x)^2/2t}\,a\Delta x.$

Proof.

Immediate.

IV.12. *Lemma*

Suppose $\{a_\nu\}_{\nu\in{}^*\mathbf{N}}\{b_\nu\}_{\nu\in{}^*\mathbf{N}}$ are internal sequences in $^*[0, \infty)$ and

$$\forall \nu \in {}^*\mathbf{N} \qquad a_\nu \doteq b_\nu.$$

Then

$$\sum_\nu a_\nu \doteq \sum_\nu b_\nu.$$

Proof.

Let $l = \inf\left\{\dfrac{a_\nu}{b_\nu}\right\}$ $\quad u = \sup\left\{\dfrac{a_\nu}{b_\nu}\right\}$.

Notice $l \approx 1 \quad u \approx 1$

$$l \leq a_\nu/b_\nu \leq u.$$

So

$$lb_\nu \leq a_\nu \leq ub_\nu$$

and

$$l\Sigma b_\nu \leq \Sigma a_\nu \leq u\Sigma b_\nu.$$

So

$$l \leq \Sigma a_\nu/\Sigma b_\nu \leq u,$$

which completes the proof.

IV.13. *Corollary*

Suppose $F(x, 0) = 0$ for infinite x and

$$F(x, t + \Delta t) = \tfrac{1}{2}[F(x + \Delta x, t) + F(x - \Delta x, t)] \tag{IV.2}$$

and $t = 2n\Delta t \not\approx 0$. Then

$$\begin{aligned} F(x, t) &\doteq \sum_{k=-\infty}^{\infty} \frac{\sqrt{2}}{\sqrt{\pi t}}\, e^{-(2k\Delta x)^2/2t} F(x + 2k\Delta x, 0)\Delta x \\ &\doteq \sum_{k=-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}}\, e^{-(2k\Delta x)^2/2t} F(x + 2k\Delta x, 0)(2\Delta x). \end{aligned} \tag{IV.3}$$

(Note: This is the analog of the usual standard result

$$\phi(x, t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}}\, e^{-(s-x)^2/2t} \phi_0(s)\, ds$$

where ϕ_0 is the initial probability density function.)

Proof.

Immediately from IV.11, IV.12, and linearity.

REFERENCES

1. Tom M. Apostol, *Mathematical Analysis: A Modern Approach to Advanced Calculus*. Addison-Wesley, Reading, MA (1957).
2. Martin Davis, *Applied Nonstandard Analysis*. John Wiley and Sons, New York (1977).
3. Abraham Robinson, *Non-standard Analysis*, Revised Edition. North-Holland, Amsterdam (1974).
4. K. D. Stroyan and W. A. J. Luxemburg, *Introduction to the Theory of Infinitesimals*. Academic Press, New York (1976).